

University of Nebraska - Lincoln

DigitalCommons@University of Nebraska - Lincoln

Faculty Publications from the Department of
Electrical and Computer Engineering

Electrical & Computer Engineering, Department
of

2-1992

Model-based target recognition using laser radar imagery

Robert Y. Li

Follow this and additional works at: <https://digitalcommons.unl.edu/electricalengineeringfacpub>



Part of the [Computer Engineering Commons](#), and the [Electrical and Computer Engineering Commons](#)

This Article is brought to you for free and open access by the Electrical & Computer Engineering, Department of at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Faculty Publications from the Department of Electrical and Computer Engineering by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.

Model-based target recognition using laser radar imagery

Robert Y. Li

University of Nebraska
Department of Electrical Engineering
Lincoln, Nebraska 68588

Abstract. Autonomous target recognition can be assisted by using CO₂ laser radar data. Of these data, range data provide 3-D geometric information and Doppler data detect boundaries of moving targets. A powerful 3-D feature extraction algorithm based on the Hough transform is used to obtain the orientations and dimensions of the target. This information is then utilized by an inference procedure that recognizes targets based on the available evidence from the sensory data. The experimental results using actual laser radar imagery are successful and the procedure can be used for the future development of a model-based expert system for target recognition.

Subject terms: target recognition; range image; Hough transform; expert system.

Optical Engineering 31(2), 322-327 (February 1992).

1 Introduction

Autonomous detection, classification, and location of objects or targets from the sensed image are some of the goals of computer vision research.^{1,2} CO₂ laser imaging sensors have been explored as useful tools for reaching these goals, because they offer a unique 3-D target sensing capability.^{3,4} Our efforts are focused on extracting meaningful features from the laser radar images and performing a classification study using expert system concepts.

For this research, the multisensor imagery was obtained from a multimodal CO₂ laser radar that has the following data types: relative range, Doppler, IR, visual, carrier intensity, and vibration.⁵ Of these data, Doppler data provide very reliable detection discriminants for mobile targets in a battlefield environment. For Doppler imagery, the ground plane is usually recorded as zero velocity. An important application of Doppler data is the detection of moving targets that are largely obscured by foliage or clutter. On the other hand, relative range data are measured by impressing a sinusoidal pattern of amplitude on the transmitted beam and by detecting the phase of the pattern in the received radiation. Range data are capable of reconstructing the 3-D geometry of the scene viewed from the sensor.⁶⁻⁹ In a real-time operation, use of fast and reliable algorithms to extract these geometric features is important. The Hough transform, used extensively here, is a technique that offers speed and accuracy for extracting 3-D planar features. The next step is to use the structure of an expert system for classifying the object. An expert system is essentially a hypothesis management system integrating these features into a decision process. By using this system, the hypotheses about the nature of the target are tested and the evidence analyzed to support or discredit those hypotheses. The dataset used in this study was relatively small, but adequate in demonstrating the concepts of model-based target recognition as envisioned here. The overall purpose was to study the methodology of building a model-based target recognition system and to develop the tools needed for building such a system.

The approach used in this study is summarized in the following three major sections: (1) data preprocessing: use of Doppler imagery to convert relative range images into absolute range images and to generate XYZ maps by taking the motion effect into account, (2) geometric feature extraction: segmentation of major vertical and horizontal surfaces by computing displacements among adjacent pixels and extraction of surface orientations and dimensions using the Hough transform, and (3) expert system approach: An inference procedure was built so that the hypothesis of the model can be evaluated based on the evidence supplied by the extracted features. A LISP program was written to implement the procedure. These three tasks will be described in more detail in the following sections.

2 Data Preprocessing

2.1 Absolute Range Conversion

The laser image dataset includes relative range images with 15-m range ambiguity. Because most targets of interest are smaller than 15 m, the relative range data cover most targets in one ambiguity level. However, because of low oblique angles, the ground plane around the target is usually covered by three to four ambiguity levels. To concentrate the analysis on the target, our first step is to segment a section of the relative range image around the target as well as the corresponding part of the Doppler image. Then, we convert this subimage from relative range to absolute range by removing the ambiguity levels.

Removal of the ambiguity levels of the relative range images is a tricky problem. The difficulty has to do with the target obstructing the ground behind it. We begin the conversion by first median-filtering the range and the Doppler data to remove noise using a 3×3 window around each pixel. For cases where the window contains range data from two different ambiguity levels, 256 is added to the low values before sorting the nine intensity values in the window. After the median-filtering, the relative range values are examined by column from the bottom of the image to the top. The crossing of ambiguity levels is detected by noting large differences between two adjacent pixels. The range value of each pixel is then adjusted accordingly. Be-

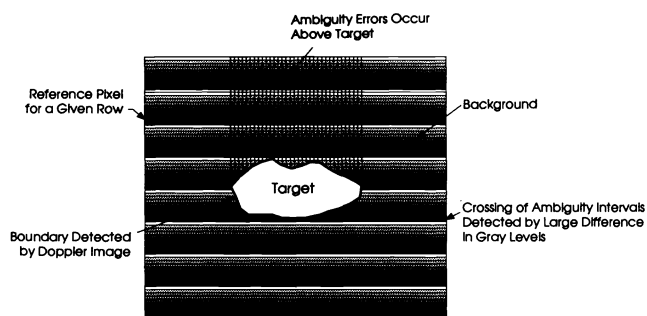


Fig. 1 Removing ambiguity error when converting a relative range image to an absolute range image in a bottom to top approach by column.

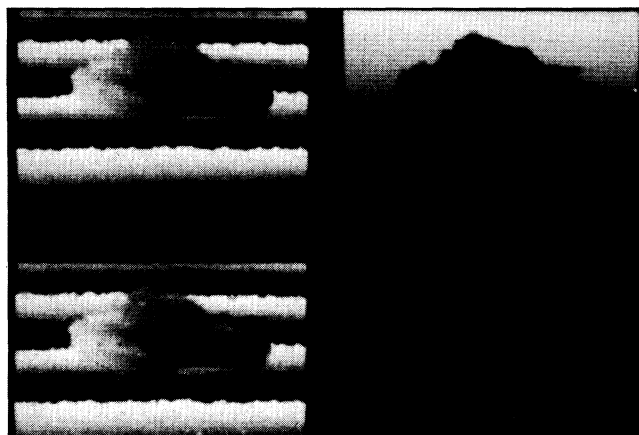


Fig. 2 Tank scene, clockwise from left bottom: original relative range image, median-filtered relative range image, converted absolute range image, and median-filtered Doppler image.

cause a target obstructs the ground behind it, this conversion from relative range to absolute range often results in ambiguity errors in the image columns above the target. This is corrected by first using the Doppler image as a mask to determine the boundary between the target and the erroneous pixels in a given column. Then in a left to right search, the columns that are in error are detected and corrected using the leftmost pixel of the search as a reference value. The concepts of the steps mentioned above are illustrated in Fig. 1.

Figure 2 shows a tank in a relative range image, the filtered relative range image, the filtered Doppler mask, and the converted absolute range image. Comparison of the original range image and the median-filtered range image shows that the salt-and-pepper noise no longer exists. The converted absolute range image shows that the range values change gradually and smoothly from pixel to pixel, and the range ambiguity is removed. The use of the Doppler channel is very effective and unique for doing the range conversion. Many current sensors do not produce ambiguous range images, but sometimes the need for processing ambiguous images still exists.

2.2 XYZ Map Generation

The range data should be converted to a sensor-based Cartesian coordinate system (x, y, z) using the known sensor gim-

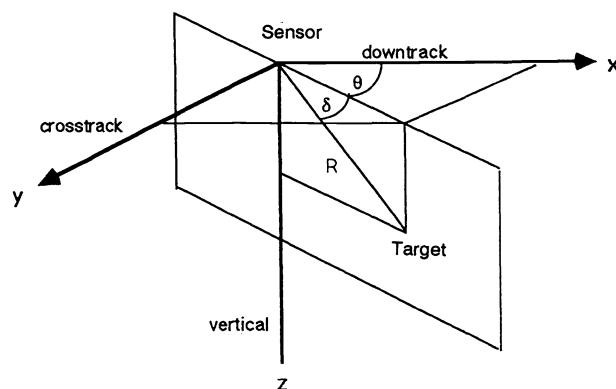


Fig. 3 Conversion from R to (X, Y, Z) ; outside gimbal is azimuth.

bal angles. The coordinate transformation can be described by the following equations:

$$\begin{aligned} \text{downtrack, } x &= R \cos \delta \cos \theta \\ \text{crosstrack, } y &= R \cos \delta \sin \theta \\ \text{vertical, } z &= R \sin \delta \end{aligned} \quad (1)$$

where δ is the depression angle and θ is the azimuth angle.

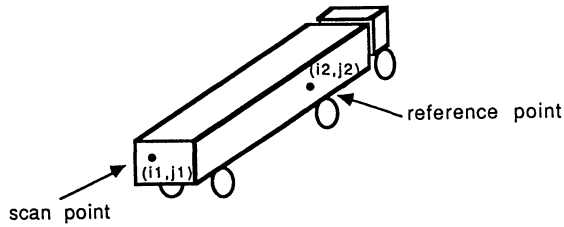
Figure 3 shows the geometry for the above coordinate transformation. The distance and depression angle from the sensor to the center of the image frame and that of the major target are available from the file's header information. The angular resolution is also available and is usually either 0.29 or 0.19 mrad. They are used in the calculation of the xyz values.

The CO_2 laser radar scans vertically bottom to top with the first scan on the left and the last scan on the right at 10,000 pixels per second. For the 256×256 images, each frame takes about 6.5 s to record. When targets are moving, the range measurement of the target is either elongated or shortened due to the time gap between recordings of different points on the same target. For example, a target moving at 5 mph moves 2.2 m in 1 s and can cause considerable range discrepancy. For most cases, the Doppler data showed that the target was moving at a very slow velocity. However, for some other cases, the target's velocity was at 5 mph or more. We decided to include a correction routine for the downtrack direction because velocity in that direction can be calculated from the Doppler data. The Doppler data are quantized in 0.09752 m/s per level. The correction assumes that the target is moving at a constant velocity based on the Doppler value of each pixel. The time gap Δt between a pixel and the reference point is calculated by dividing the number of pixels scanned between the two points by 10,000 pixels/s. The adjustment in downtrack direction is computed as

$$\Delta x = V_x \Delta t = 0.09752 (D_s - 120) \Delta t \quad (2)$$

where D_s is the Doppler value at the scan point, and the reference Doppler value for a motionless pixel is equal to 120. We adjust every pixel on the moving target by $\pm \Delta x$. See Fig. 4 for an illustration of motion correction on downtrack distance using the Doppler index.

256 x 256 Image



$$\Delta t \text{ (time elapsed)} = (\# \text{ of pixels from ref. point}) \times (6.5/256/256)$$

$$\Delta x = V_x \times \Delta t = 0.09752 \text{ m/s} \times (D_s - 120) \times \Delta t$$

Fig. 4 Motion correction on downtrack distance using Doppler index. Every pixel on the moving target is adjusted.

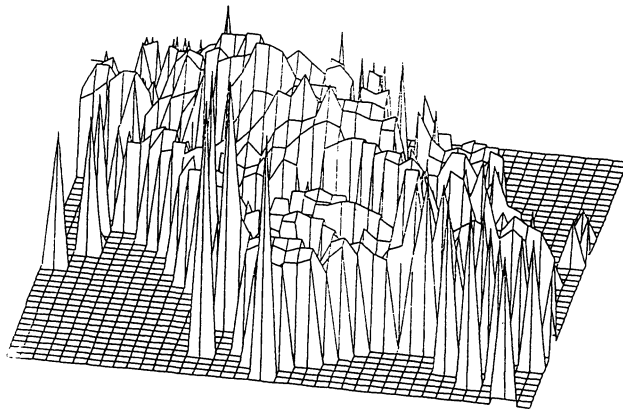


Fig. 5 Perspective view of a tank scene.

The xyz values generated in this step are subsequently used for surface extraction and Hough transformation calculations. They are stored in an ASCII file for later use. To see that the xyz values are basically correct, we plotted the perspective view of a target using the xyz values with MATLAB (a matrix-based mathematics software). Figure 5 shows the plot of an M-60 tank with its pixels extracted using a Doppler mask. Although some noise still exists, note that the lower section, made up of the hull and the tracks, is box-like. The upper section shows the existence of a turret with part of the gun barrel in front.

3 Geometric Feature Extraction

3.1 Surface Extraction

Two very useful preprocessing operations were used in surface extraction. The first tests for vertical displacement between two adjacent downrange pixels in order to separate vertical pixels from horizontal pixels. The second separates these candidate vertical pixels into left- and right-facing vertical pixels.

The vertical test applies a 3×3 convolution mask to the Z components of all the pixels in the image. The pixels are treated as a spatial array in which each row is at a constant

depression angle and each column is at a constant azimuth angle. The convolution operation is as follows:

$$\mathbf{SLP} = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ -1 & -1 & -1 \end{bmatrix} (\mathbf{Z} \text{ value}) . \quad (3)$$

If the convolution is applied to row n , the difference in Z value between row $n+1$ (a smaller depression angle) and row $n-1$ (a larger depression angle) will be positive for any rising slope. The \mathbf{SLP} value is normalized over distance and multiplied by a scaling factor. The result of this convolution at each pixel is then compared to a set of threshold values to isolate horizontal surfaces and vertical surfaces.

The second operation classifies the resulting vertical surfaces as left- or right-facing surfaces. The operation uses another 3×3 convolution mask that is applied to the X component (downtrack) of all of the vertical pixels derived from the first operation. The second convolution mask is given by

$$\mathbf{F} = \begin{bmatrix} 1 & 0 & -1 \\ 1 & 0 & -1 \\ 1 & 0 & -1 \end{bmatrix} (\mathbf{X} \text{ value}) . \quad (4)$$

The results of applying this mask at column n of the image will be positive for a left-facing surface since the downtrack value at column $n-1$ (to the left of n) is larger than the downtrack values at column $n+1$ that is to the right of column n . Similarly, the results for right-facing surfaces will be negative. For surfaces that are perpendicular to the sensor's line of sight, \mathbf{F} will be close to zero. Using this information, the left- and right-facing vertical pixels can be distinguished.

3.2 Three-Dimensional Hough Transform

In terms of extracting information from 3-D surfaces, many approaches based on the techniques of differential geometry exist. However, these approaches have mainly been used on synthetic or short-range laboratory imagery. They will not work well for data with even a small amount of noise. For long-range data, such as the data used here, the Hough transform can be used to extract geometric information from potential targets. The Hough transform is usually used for curve or line detection in a 2-D intensity image.^{10,11} Because each pixel is individually transformed, the Hough technique is relatively unaffected by gaps and noise in the curve. The idea can be extended to 3-D data to detect and measure planar surfaces.^{12,13}

Ordinarily a plane is described by the following equation:

$$Ax + By + Cz = 1 . \quad (5)$$

However, we want to relate coefficients A , B , and C to the pan angle (θ), tilt angle (δ), and normal distance (D). As a result, a plane in the 3-D space can be described by the following equation:

$$\cos\delta \cos\theta x + \cos\delta \sin\theta y + \sin\delta z = D , \quad (6)$$

where θ , δ , and D are pan angle, tilt angle, and normal distance, respectively.

To detect a plane, we check for any peak in the accumulator of these three parameters (θ, δ, D). When θ , δ , and D extend over a large range, the calculation can be very time-consuming. However, since most surfaces are either vertical or horizontal, the Hough transform for 3-D space can be simplified. Because the tilt angle of a vertical surface is zero, the equation reduces to

$$\cos\theta x + \sin\theta y = D \quad (7)$$

For horizontal surfaces, because the tilt angle is 90 deg, the equation reduces to

$$z = D \quad (8)$$

These simplifications reduce calculation time greatly, yet still produce reasonably good results.

Our first goal is applying the Hough transform to the data is to estimate target orientation by detecting a dominating vertical surface in the target. This is done by locating the peak in the (θ, D) accumulators using the xy values of each pixel. We initially included all the pixels of the image in calculating votes for each accumulator. Two target types, the M-60 tank and the fuel truck, were used. Because neither the M-60 tank nor the fuel truck consisted of large vertical flat surfaces, we obtained poor results. We then tried to use only the target pixels extracted with the Doppler mask. Still, no clear peak was found to give us the correct target orientation. We then reasoned that most vehicles had either a flat top or a large deck and the portion below that flat horizontal surface was more box-like. Therefore, vertical pixels below the major flat surface are more suitable for the calculation of target orientation. Using this approach, correct orientation angles were obtained for most targets under consideration.

3.3 Feature Extraction

After target orientation is determined, we can calculate geometric features such as height, width, and length of the target and the percentage of pixels above the target deck. These features are used as the inputs to a model-based inference net. These dimensions were computed by counting the valid intervals within the Hough space along the proper orientation. The specific steps involved in estimating the orientation and the dimensions of a target are summarized below

1. With the Doppler data as a mask, segment the target's major horizontal surface using the Hough transform.
2. Using only vertical pixels below the major horizontal surface and only angles from -90 to 90 deg in the Hough space, find the target orientation.
3. With the angles fixed, find the target dimensions by counting valid intervals of the normal distance along these angles.
4. Calculate "percentage of pixels" above and below the major horizontal surface (the deck in the case of the tank).

To find the feature of percentage of pixels, we first use the Doppler channel to segment the target from the background. Then we look for a major horizontal surface by using the Hough transform. Using the major horizontal sur-

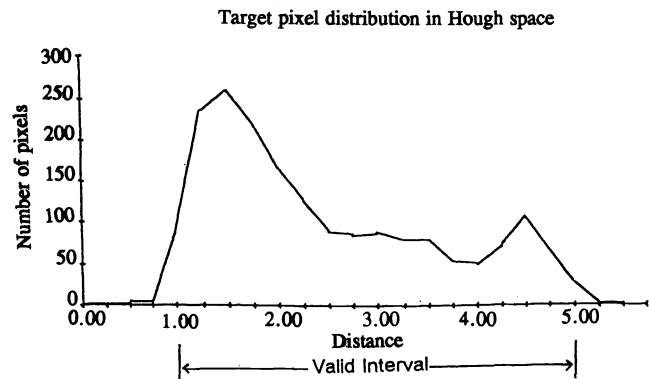


Fig. 6 Measuring dimensions by estimating valid intervals of a Hough space. The measurement here is for the width of a tank.

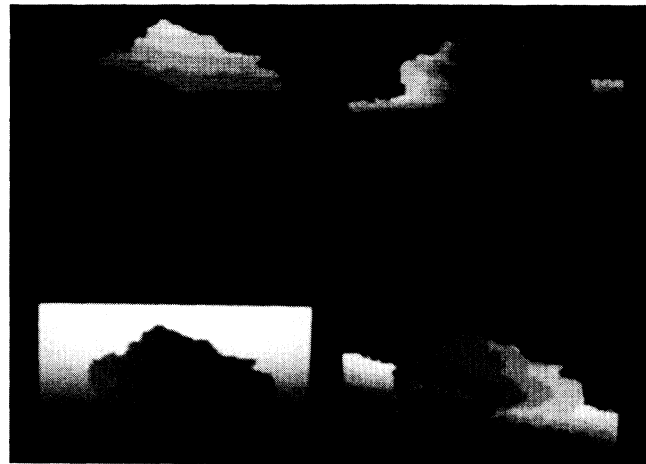


Fig. 7 Hough transform results shown here by layers of different intensities representing pixels in different accumulators. Clockwise from left bottom: original range image, horizontal intervals, left-facing vertical intervals, and right-facing vertical intervals.

face as a reference, one can easily count the number of pixels above the surface as a percentage of the total number of pixels within the target. As long as the Doppler data are available, there should be no difficulty in obtaining the percentage of pixels. If the Doppler data are not available, obtaining a clean segmentation of the target from the ground by some other means is necessary.

To measure dimensions, we use the Hough transform to count distance intervals in the horizontal direction, left-facing vertical direction, and right-facing vertical direction. Figure 6 shows a Hough-space histogram indicating the valid interval for the width measurement of a tank. By truncating both ends of the curve, we obtain an approximation of the dimension. Figure 7 shows the projected Hough space of an M-60 tank being intensity-sliced in the horizontal direction, the right-facing vertical direction, and the left-facing vertical direction. The left-facing intervals are at a 60 deg norm, and the right-facing intervals are at a -30 deg norm, and each interval is 1 m wide. The multiple layers displayed by different intensities assigned to the intervals show how the pixels are allocated in the Hough accumulators. For

Table 1 Extracted target features.

Target Type	tank ¹	tank	tank	tank	truck ²	truck
Analysis Result						
Angle, deg ³	60	10	10	-70	20	-70
Distance, meter	314	307	880	300	300	371
Width, meter	4.0	5.0	3.75	3.5	3.75	3.0
Length, meter	6.5	5.25	5.0	6.0	12.0	12.0
Height, meter	3.0	3.5	3.5	3.0	3.25	3.25
% pixel above deck	29.8	46.1	30.5	33.3	6.4	3.8
Total no. of target pixels	1940	1365	466	1255	1551	1448

Note: 1. M-60 tank is 3.6 meter wide, 6.9 meter long and 3.2 meter high
 2. Fuel truck is 2.46 meter wide, 12 meter long and 2.64 meter high
 3. Perpendicular angle to the dominating vertical surface extracted by the Hough transform

**Fig. 8** Projection of six different targets to the horizontal plane. The four to the left are tanks, and the two to the right are fuel trucks.

illustrative purposes, 1-m-wide intervals were used in the figure to limit the number of layers displayed. In calculating actual target dimensions, 0.25-m intervals are used.

As a summary, Table 1 lists the results of computing target orientation and dimensions for six different targets involving four tanks and two fuel trucks. Figure 8 shows the top-down views of these targets in a two-dimensional xy plane. The Hough transform results appear quite consistent with the angular orientations of these top-down views. Those measurements on height, length, width, and the percentage of pixels above the deck are used as evidential inputs to the expert system that is described next. Note that estimated measurements do not agree exactly with the true specifications. This is due to the approximation method mentioned above as well as the noise inherent in the range data.

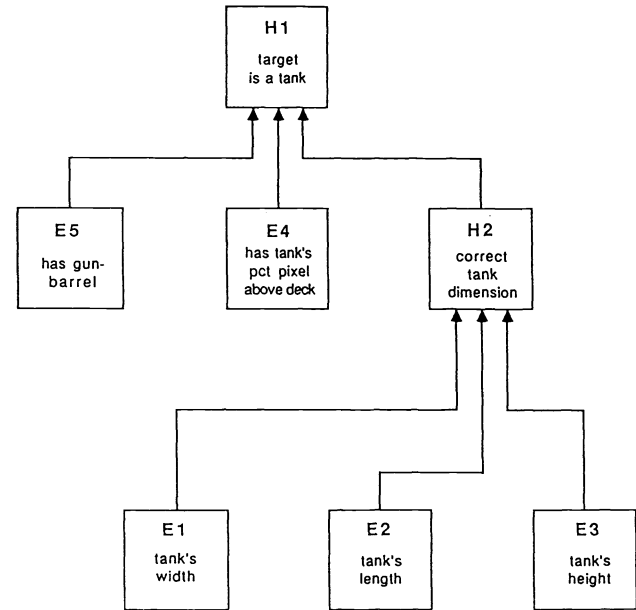
4 Expert System Approach

Our image analysis of the range data produces certain descriptions of the target, such as the target dimensions and the percentage of pixels above the deck. These descriptions are then processed by a model-based inference procedure in the context of an expert system. Figure 9 shows such an inference net for the hypothesis that "the target is a tank." The recognition model is essentially a rule-based decision network. Given certain assumptions, *a priori* probabilities, and conditional probabilities, the *a posteriori* probability is then computed for each node based on Bayes's rule outlined below

$$P(H/E) = P(E/H) * P(H)/P(E) \quad (9)$$

where

$$P(H) = \text{a priori probability that hypothesis } H \text{ is true}$$

**Fig. 9** Inference net for identifying a tank.

$P(E)$ = *a priori* probability of the evidence

$P(E/H)$ = conditional probability that if hypothesis H is true then evidence E exists

$P(H/E)$ = *a posteriori* probability that H is true in light of evidence E .

Using Fig. 9 as an example, H_2 is the hypothesis that the target's overall dimensions are correct, and E_2 is the evidence that the target's length is correct. If we estimate that the frequency of obtaining the correct overall dimension is three out of 10 in a given scenario, then $P(H_2) = 0.3$. Normally, $P(E_2/H_2)$ is 1. If $P(E_2)$ is available, then the *a posteriori* probability $P(H_2/E_2)$ can be computed by using Bayes's rule. In the model-building process, the *a priori* probabilities associated with each node must be first known or estimated. Then LS and LN , where LS equals measure of support for the hypothesis if E is present and LN equals measure of discredit for the hypothesis if E is missing, are calculated for each node. Essentially, $LS = P(H/E)$ and $LN = P(\sim H/\sim E)$, where $\sim H$ means that H is not true.

New information is added to the certainty factors of the evidence nodes and $P(E)$ is adjusted using the following equations:

$$\begin{aligned} P'(E) &= P(E) + [1 - P(E)] \cdot CF/5, & CF > 0, \\ &= P(E) + P(E) \cdot CF/5, & CF \leq 0. \end{aligned} \quad (10)$$

These changes in $p(E_2)$ yield new values for $p(H_2)$, which are calculated by using Bayes's rule and the values of LS and LN . In this example, H_2 is the hypothesis that "all three measured dimensions are correct for the target." These probabilities are propagated up the inference network using an AND relationship such that the new probability of H_2 is determined by the minimum value obtained from the three-dimensional measurements. This probability propagation is performed throughout the entire system. Finally, the probability at the root is calculated, and a CF value is generated

for the hypothesis H_1 . The approach taken here is very similar to the hypothesis testing procedures developed in the PROSPECTOR expert system.¹⁴

The inference procedure is implemented in LISP running on an IBM AT compatible PC. The model has been extended to accommodate multiple hypotheses. For instance, after installing several top-level hypotheses, the program can estimate the probabilities of each hypothesis and choose the one most likely to be true. We tested the program with our image analysis results in Table 1 using a battlefield scenario in which equal numbers of tanks, APCs, 2½-ton trucks, and fuel trucks exist. Their dimensions are available from the Jane's report. The results showed that all tanks except one were identified as tanks and both fuel trucks were identified as fuel trucks. In all the cases, both the length and the percentage of pixels were the major features that separated the vehicle types.

5 Summary

The purpose of this paper is to describe the procedures of utilizing laser radar imagery to classify target vehicles. For range data analysis, our approach to preprocessing and feature extraction is straightforward and effective. Doppler imagery is very useful because of its capability of extracting the silhouette from the target. The geometric information obtained by the 3-D Hough transform is essential to the successful operation of a model-based recognition procedure. Our experimental results showed that by using appropriate range and Doppler measurements, a model-based expert system can perform the functions of recognizing potential targets in a 3-D environment.

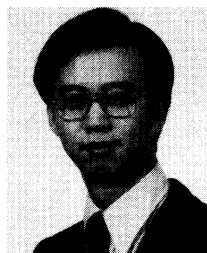
For future research, the usefulness of infrared and visible data should be explored. Range data alone produce good detection results for most scenarios, but the passive data can provide additional features that will help reject decoys and natural objects with dimensions similar to those of the targets.^{15,16} A synergistic approach involving both active and passive data should increase the accuracy of a recognition system.

Acknowledgments

This research was supported by the Air Force Office of Scientific Research under contract F49620-88-C-0053. The author would also like to thank one anonymous reviewer for his valuable suggestions.

References

1. B. Bhanu, "Automatic target recognition: state of the art survey," *IEEE Trans. Aerospace Electron. Sys.* **AES-22**(4), 364-379 (1986).
2. A. K. Jain and R. Hoffman, "Evidence-based recognition of 3-D objects," *IEEE Trans. Pat. Anal. Mach. Intell.* **PAMI-10**(6), 783-802 (1988).
3. R. A. Jarvis, "A perspective on range finding techniques for computer vision," *IEEE Trans. Patt. Anal. Mach. Intell.* **PAMI-5**(2), 122-139 (1983).
4. D. Nitzan, A. E. Brain, and R. O. Duda, "The measurement and use of registered reflectance and range data in scene analysis," *Proc. IEEE* **65**(2), 206-220 (1977).
5. G. S. Bain and W. K. Cunningham, "Multifunction laser radar," Technical Report, LTV Aerospace and Defense Co. (May 1986).
6. R. S. Loe, J. G. Landowski, and C. M. Bjorklund, "Feature extraction for undersampled objects in range imagery," SPIE 1988 Technical Symposium Southeast on Optics, Electro-Optics and Sensors, Orlando, Florida, April 4-8, 1988.
7. P. Besl and R. Jain, "Range image understanding," in *Proceedings of Conference of Computer Vision and Pattern Recognition*, IEEE Computer Society Press, pp. 430-449 (1985).
8. R. Y. Li, "Range image processing for object recognition," in *Proceedings of the 30th Midwest Symposium on Circuits and System*, Elsevier Science Publishing Co., pp. 1226-1229 (1988).
9. D. T. Kuan and R. J. Drazovich, "Model-based interpretation of range imagery," *AAAI Proc. Nat. Conf. on AI*, pp. 210-215 (1984).
10. A. P. Reeves and W. Lou, "Object recognition using the tuned generalized Hough transform," *IEEE Syst., Man, Cyber.*, pp. 124-129 (1987).
11. T. Shibata and W. Frei, "Hough transform for target detection and infrared imagery," in *Technology and Applications of Image Understanding*, *Proc. SPIE* **281**, 105-109.
12. R. Y. Li, C. M. Bjorklund, and R. S. Loe, "Matching of planar surfaces for 3-D terminal homing," in *Proceedings of NAECON Conference*, IEEE (1983).
13. Milgram and C. M. Bjorklund, "Range image processing: planar surface extraction," in *IEEE Proc. 5th Int. Conf. Pat. Rec.*, pp. 912-919 (1980).
14. A. Waterman, *A Guide to Expert Systems*, pp. 55-60, Addison Wesley Publishing Co. (1985).
15. H. W. Rose and J. C. Rachal, "Three dimensional automatic target recognition," in *Technology and Applications of Image Understanding*, *Proc. SPIE* **281**, 116-125 (1981).
16. S. K. Rogers, C. W. Tong, M. Kabrisky, and J. P. Mills, "Multi-sensor fusion of lidar and passive infrared imagery for target segmentation," *Opt. Eng.* **28**(8), 881-886 (1989).



Robert Y. Li received his BS degree from Duke University in 1972, MS degree from Purdue University in 1976, and PhD degree from the University of Kansas in 1981, all in electrical engineering. For three years, he was associated with IBM, working on digital system design. In 1981, he joined Lockheed Palo Alto Research Laboratory as a research scientist, working on projects in image-based missile guidance and 3-D target recognition. He also taught graduate

courses at the University of Santa Clara. Since August 1986, he has been an assistant professor of electrical engineering at the University of Nebraska in Lincoln. His primary areas of interest are in image processing and artificial intelligence, and he has published extensively in these areas.